SHIVAM LALAKIYA

Boston, MA | 781-363-8482 | lalakiya.s@northeastern.edu | LinkedIn | portfolio

EDUCATION

Northeastern University, College of Engineering | Boston, MA — **June 2023 (Expected)**
Master of Science in Data Analytics Engineering — GPA: 3.9/4.0
Relevant Courses: Foundation of Data Analytics, Data Management, Comp. and Visualization, ML, NLP, Data Mining, Cloud Computing

Sardar Vallabhbhai National Institute of Technology | Surat, India — **May 2021**
Bachelor of Technology in Electronics and Communication Engineering — CGPA: 8.10/10.0

TECHNICAL SKILLS

Languages/ Frameworks: Python, R, SQL, C/C++, Shell scripting, MATLAB, NoSQL, Java, Scala
Tools and Technologies: Machine Learning, NLP, Computer Vision, Reinforcement Learning, CNNs, RNNs, Power BI, Tableau, ETL, ML-flow, Airflow, Pyro, AWS, git, CI/CD, Snowflake, Kafka, Spark, Hadoop, Hive, Docker, Kubernetes
Soft Skills: Leadership, Management, Communication, Public Speaking, Teaching, Content Writing

EXPERIENCE

Ring Therapeutics | Data Science co-op | Boston, MA — **May 2022 – Dec 2022**
- Provided analytical insights regarding binding sites, tissue, and cell specificity from sequenced data collected by the Discovery team from the patients' protein sequences to enhance the virology team's outcome and develop Anello-backed programmable medicine.
- Leveraged NLP-based models for genes/protein sequences and implemented positional encoded models and Graph Neural Networks for predicting tropism to create viral vectors that can safely and effectively deliver therapies to target cells and tissues.
- Built ETL Docker containers and pipeline to load data into fasta format using Airflow DAG to eliminate manual data loading to the server.
- Created interactive front-end for company-wide utilization of developed ML models and functionalities.

IIT Madras | Research Assistant | Chennai, India — **Jan 2020 - Jan 2021**
- Executed a 'Caching with Deep Learning' project, leveraging RNNs and LSTM models for time-series prediction of user preferences, resulting in a 90% success rate in accurately forecasting future requests.
- Preprocessed and meticulously filtered a substantial 12GB dataset using Pandas, employing advanced data manipulation techniques to ensure optimal training of deep learning models.
- Designed and implemented an LSTM-based caching policy, outperforming traditional policies (LIFO, LRU, and LFU) with a 130% improvement in hit rates, as demonstrated through extensive benchmarking and performance analysis.

Sense Grass | Data Science Intern | Bengaluru, India — **July 2019 - Dec 2019**
- Achieved an impressive 90% accuracy rate in predicting crop diseases and yields, resulting in enhanced decision-making capabilities and improved agricultural outcomes.
- Developed and evaluated machine learning models using satellite images and vegetation indices, leading to precise crop yield predictions and accurate identification of diseases.
- Implemented models resulting in a 25% reduction in crop disease-related losses and a 15% increase in crop yield, optimizing resource management and enhancing farm profitability.

Sardar Vallabhbhai National Institute of Technology | Research Assistant | Surat, India — **July 2018 – Dec 2018**
- Developed a "Multimodal Biometric System" project combining Iris, Facial, Speech Recognition, and fingerprints using Convolutional Neural Networks (CNNs), achieving an 85% precision rate on multiple datasets.
- Deployed the model on cloud-based platforms (AWS and GCP) for real-time biometric authentication, achieving a reliable and accurate system with an F1 score of 0.87 and a precision of 0.91.
- Applied the system to automate attendance in the department, saving 15 minutes per class and enhancing security in the professor's cabin.

ACADEMIC PROJECTS

Stock Market sentiment and time series analysis (Python, LSTM, Stream lit, Web-scrapping, Java) — **May 2022 – Aug 2022**
- Extracted recent 30 days' stock data from WSJ and historical prices from Tiingo API using ticker symbols to store them in CSV format.
- Developed the Naïve-Bayes model to predict trends and future prices using the LSTM model with more than 90% accuracy.
- Created interactive front-end utilizing Stream lit library to find insights into any listed stock and reach a mass audience.

Flight Booking Website (JavaScript, CSS, HTML, NodeJS, CI/CD, REST API, Docker) — **Sep 2021 – Dec 2021**
- Created a flight booking website that fetches live flight ticket information from sky scanner API using React.JS, NodeJS for backend infrastructure containing controllers, and MongoDB to persist data.
- Built User, Admin, and Airline modules with REST APIs for performing CRUD operations like creating and updating user, flight, and airline information and creating and applying deals to different flights for different users.
- Deployed the React framework in Netlify and the NodeJS portion in Heroku and Dockerized the application to run in a remote environment and support the CI/CD pipeline.

Analytics and Visualization using R programming (R, SQL, Tableau, ETL, Time-series forecasting, A/B testing) — **Sep 2021 – Dec 2021**
- Performed clustering, probabilistic analysis, text mining, and A/B testing to find insights on University and E-commerce datasets.
- Derived the most affecting factors for ranking the top 100 universities and reasons behind the changes in hierarchy.
- Obtained the correlation between discounts, sales, and profit for e-commerce websites and concluded that festive seasons have 35% higher sales and non-festive seasons have 30% higher deals.